GESIS Leibniz-Institut für Sozialwissenschaften

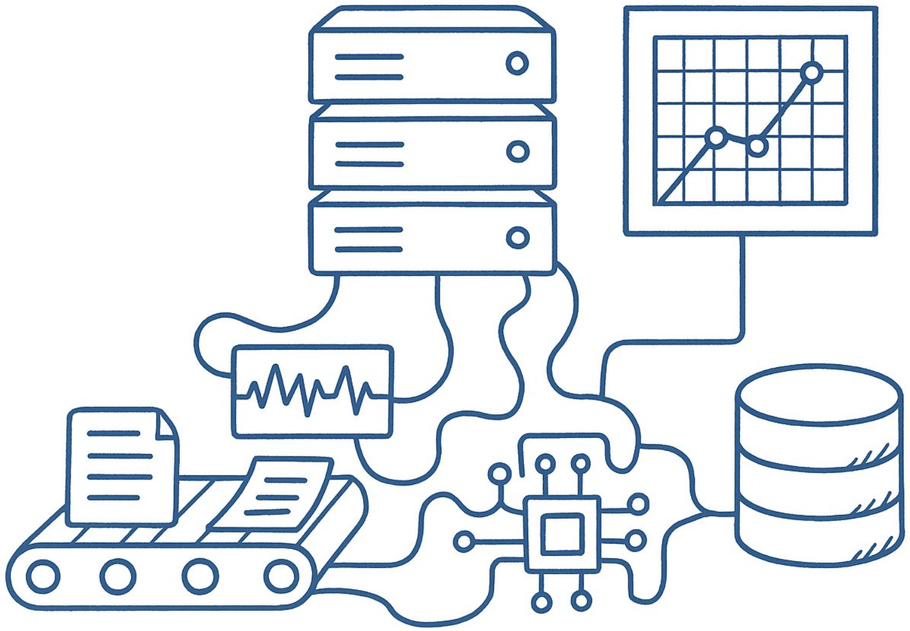
The Role of Research Infrastructure for Responsible Research with Online Platform Data

Leibniz Gemeinschaft

Katrin Weller
katrin.weller@gesis.org

Berlin, Data Access Days, 25 September 2025

What is research infrastructure?

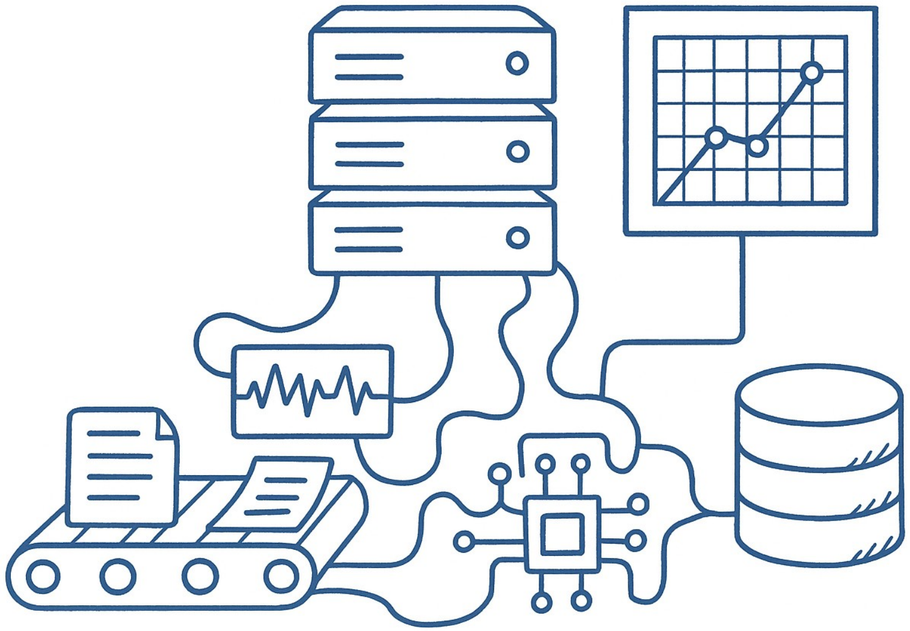


Research infrastructure can refer to both **physical** and **organisational** measures and resources that enable high-quality, reproducible, and collaborative scientific work.

Examples include:

Data and data access, facilities and hardware, tools and pipelines, consulting and best practices, organisational frameworks.

What is research infrastructure – and why do we need it?



Research infrastructure

- offers access to resources that individuals can't provide alone
- reduces individual's workload by offering centralised tools
- establishes quality standards, e.g. for reproducibility, good scientific practice
- ensures long-term sustainability
-

What is research infrastructure – for online platform data?

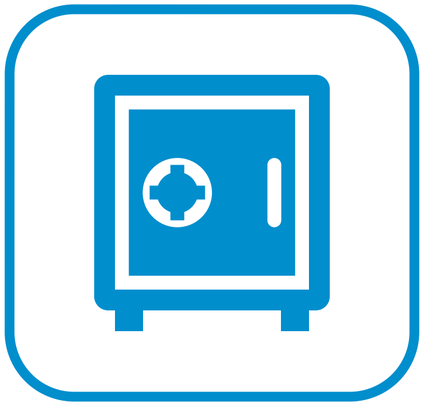
Online platform data pose new challenges and will require new solutions.

But solutions **can build on foundations** that exist across several areas, including

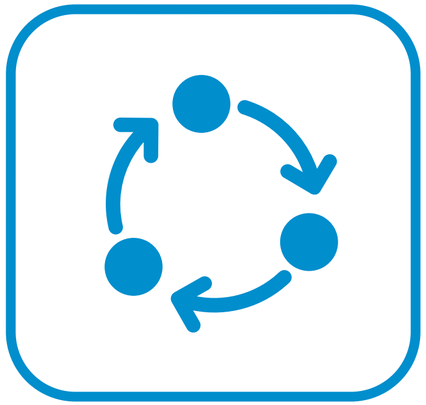
- data archiving and sharing practices across disciplines (e.g., access regimes, documentation and metadata)
- handling sensitive data (e.g., in the life sciences, health sciences, or for administrative data) in trusted research environments

3 core elements for improving research infrastructure for data access under the DSA

Researcher support



Secure Processing Environment



Secondary Data Usage

1 Researcher support

Including, for example:

- (Certified) **training** on handling sensitive data
- **Research data management plan** templates and guidance including, e.g.:
    - Data minimisation strategies
    - Deletion protocols
- Data **quality** indicators and support

→ This is important for all possible ways in which online platform data is accessed and collected (including scraping, data donations, Article 40.12).

Why it matters:

DSA Article 40(8e): their application demonstrates that their access to the data and the time frames requested are necessary for, and proportionate to, the purposes of their research. […]

DSA Delegated Act, Article 9 (f): existence of training on data security and protection of personal data received by the applicant researchers

2 Secure Processing Environment

Building a Secure Processing Environment based on the **Five Safes** framework that includes, for example:

- Researcher vetting procedures (potentially via federated login system)
- Project vetting procedures
- Researcher training
- Access rights
- Secure storage and computing
- Audit trails
- Output control
- …

https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/

Why it matters:

DSA Delegated Act Article 9 a-l, for example:
(b) relevant network security measures, encryption, access control mechanisms, backup policies, data integrity mechanisms, incident response plans;

DSA Article 40 (d):
they are capable of fulfilling the specific data security and confidentiality requirements corresponding to each request and to protect personal data, and they describe in their request the appropriate technical and organisational measures that they have put in place to this end;

Green, E., & Ritchie, F. (2023). The present and future of the Five Safes framework. Journal of Privacy and Confidentiality, 13(2), Article 2. https://doi.org/10.29012/jpc.831

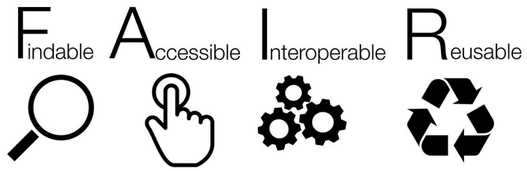
3 Secondary Data Usage

Enable data archiving and creation of reusable („common use“) datasets, this includes, for example:

- Creating reference datasets for topics of broad interest (e.g. elections, longitudinal samples)
- Depositing data resulting from reasoned requests for secondary usage and reproducibility (e.g. during publishing processes).

Goal: make data „**FAIR**“: findable, accessible, interoperable, reusable. This includes, e.g.:

- Documentation and metadata standards
- Persistent identifiers



Why it matters:

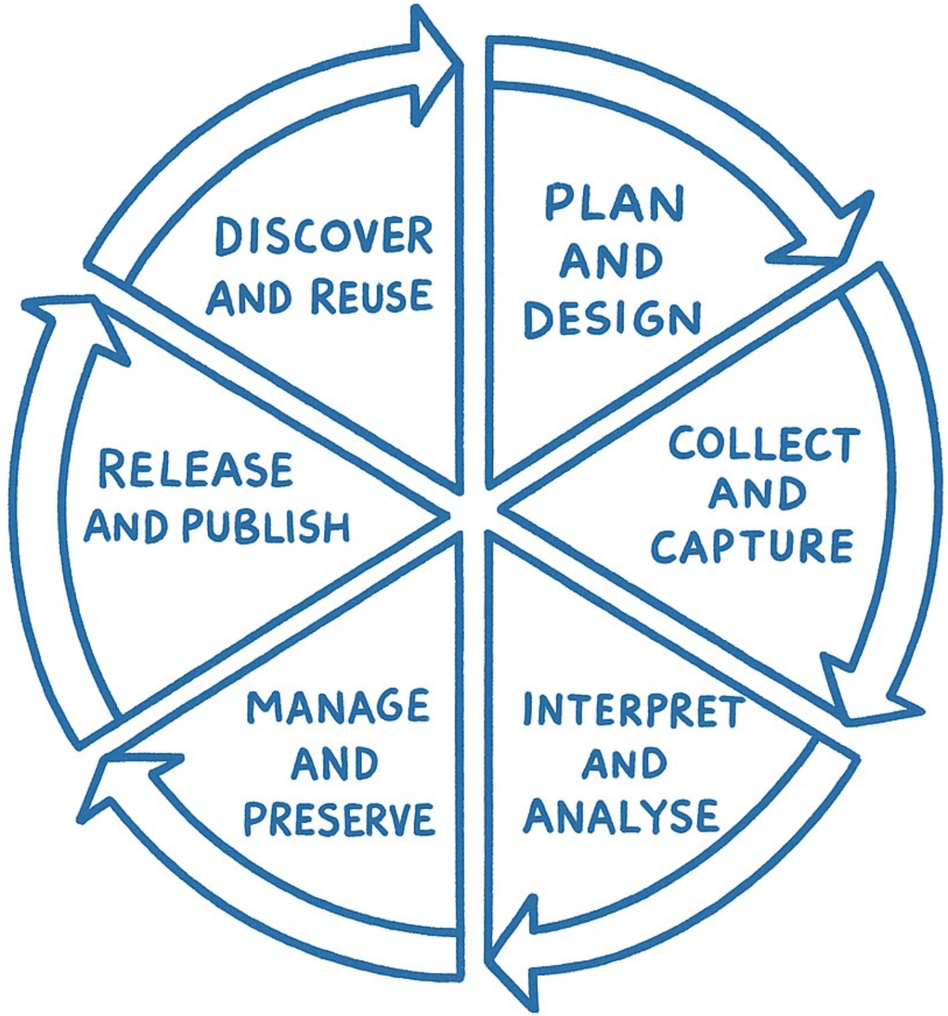
DSA Article 40 (8g):
they have committed themselves to making their research results publicly available free of charge, […]

DSA Delegated Act, Recital 13: Where the requested data are also available through other sources, the Digital Services Coordinator of establishment should assess whether the request for such data in the data access application is duly justified, having regard to the information in the data access application.

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016).

GESIS Leibniz-Institut für Sozialwissenschaften

Bigger picture: Envisioning infrastructure along a research data lifecycle



Bigger picture: Envisioning infrastructure along a research data lifecycle

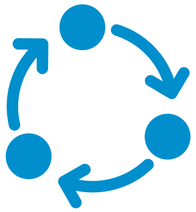
GESIS Leibniz-Institut für Sozialwissenschaften

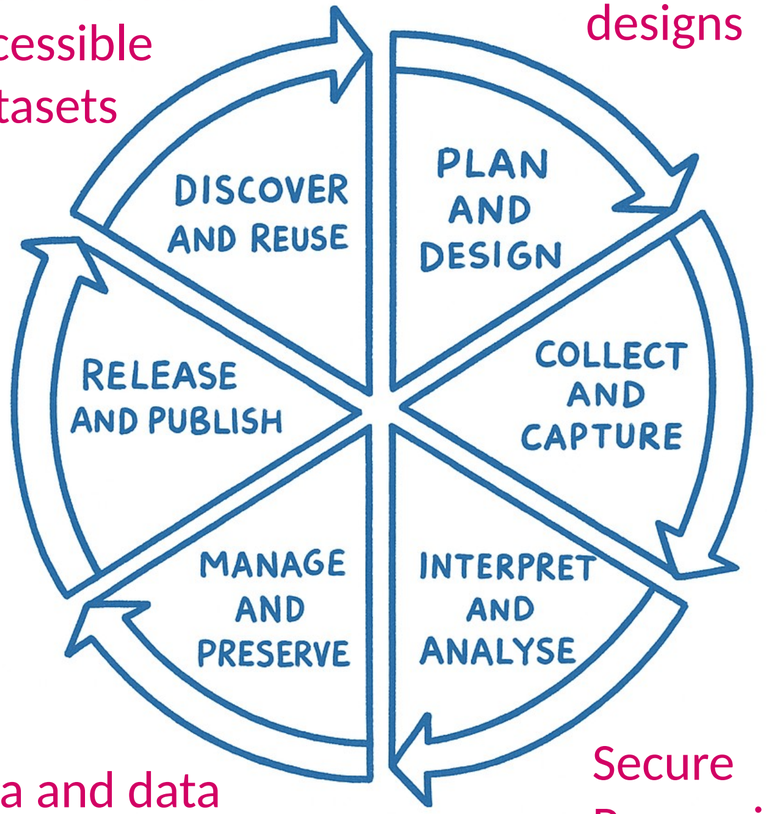
Responsible data collection designs

Archived and accessible datasets

data management plans

Reference datasets

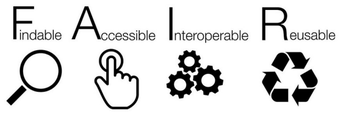




Data requests DSA 40(4) and 40(12)

Shareable data

FAIR criteria



Vetting, training, output control

Five Safes

Metadata and data documentation

Secure Processing Environments for data analysis

Are we there yet?

Currently happening at GESIS:

- Preparations for **consulting** activities for requests (including data management and data quality components)
- Re-organisation of our **Secure Processing Environment** so that it fits criteria from DSA Delegated Act and usage requirements

Interim setup available for potential joint requests under Article 40(4)

- Using our interim secure processing setup (ISO/GDPR compliant)
- Taking over responsibilities around data management

Reach out to us today/tomorrow!

Are we there yet?

“Article 40 Won’t Save Us (but it still might help)” – Rebekah Tromble

Will *infrastructure* save us?
It will help!

Next steps for the community

Getting ready for Article 40(4):

- Create first use cases of high quality, respecting the complex formal requirements → demonstrate that researchers can fulfil the high expectations on researcher responsibilities
- Jointly discuss practical challenges during the preparation of necessary materials

Connect, exchange experience, learn together!

Questions, ideas, or feedback?

I’d be happy to continue the conversation.

Katrin Weller
katrin.weller@gesis.org

Leibniz Gemeinschaft